(1)

 (a) The accuracy of the source test data when the model was trained using the source training data was: 0.86

(b) 0.78.

 It seems that a classifier that was learned in the source domain didn't generalize well in the target domain.

(c) 0.55

Which might be expected since the number of training points used is very limited and nit enough for the model to perform well.

(d) 0.78

the performance of the classifier in the target domain enhanced when it was trained with both labeled training points in the two domains, compared to the case where it was only trained using the labeled target training points.

(e)

The classifier performed relatively well when it was trained and tested on the source domain. The classifier performed poorly in the target domain when it was trained only with the limited number of labeled data in the target domain, but the performance enhanced when we added the source training points to our training procedure. We notice that the classifier performed equivalently in the target domain when it was trained using the source training only, and when it was trained using the source training with extra labeled data in the target domain, and this probably because the number of labeled points in the target domain is very few and it will not make a difference in enhancing the model's performance on the target domain.

(f) (i)

Target domain mean= [-1.01279193, -0.33115602]

Target domain covariance matrix = [[0.80377391, 0.36188612],[0.36188612 ,18.97935487]]

Source domain mean= [1.13720534, 0.99150589]

Source domain covariance matrix = [[3.83820387, 0.15370852],[0.15370852 ,0.74990653]]

(ii)

Target domain mean= [-1.05830525, -0.45760988]

Target domain covariance matrix = [[0.76377201, 0.30418792],[0.30418792 ,19.02092692]]

Source domain mean= [1.14916266, 1.09731636]

Source domain covariance matrix = [[3.70506986 -0.08005567],[-0.08005567 0.65845626]]

The second approach yields better results because the two distributions $P_s(x)$ and $P_t(x)$ are different due to the covariate shift. So, it is more likely that we will get more accurate results when the parameters on each domain are estimated separately.

(g)

When use the weighted source training points without the target training points for training, we got an accuracy of: 0.81

But when we add the few labeled target points in addition to the weighted source training points in the training process, we got an accuracy of: 0.89

(h)

The transfer learning approach performed better than the supervised learning approach because the source training points have been weighted to compensate for the change in the density function in the two domains. The supervised approach trained the classifiers using the few target training points and the original source training point without any correction factor, and hence it is expected to perform worse that the transfer learning approach.

We also notice that the best accuracy achieved was when the weighted source training data and the few labeled target training data were all utilized.